REMARKS

Applicants propose to amend claims 1 and 9, add new claim 15, and cancel claim 10. Upon entering the amendments, claims 1-4, 6-9, 11, 12, and 15 will be pending in the application.

The examiner rejected claim 1 under 35. U.S.C. § 103(a) as being unpatentable over European Patent Application EP 1,271,469 to Marasek et al., in view of U. S. Patent No. 5,796,916 to Meredith (Meredith), in view of U.S. 6,081,780 to Lumelsky (Lumelsky), in view of International Publication No. WO 02/097590 to Cameron (Cameron). But we believe that the combination of Marasek's system with the teachings of Meredith, Lumelsky, and Cameron, does not produce the invention of claim 1, as now amended. We explain this below.

Claim 1 requires:

> receiving a spoken utterance including at least one of a command to be executed by the handheld device and a name to be dialed by the handheld device;
> in response to receiving the spoken utterance:
> extracting one or more prosodic parameters from the spoken utterance;
> performing speech recognition on the spoken utterance to generate a recognized word;
> from the recognized word that is generated from the speech recognition, synthesizing a nominal word;
> generating a prosodic mimic word from the synthesized nominal word and the extracted one or more prosodic parameters…; and
> if the recognized word includes a command, executing the command on the handheld device, and if the recognized word includes a name, dialing a number associated with the name. [emphasis added]

In other words, the claim requires that a sequence of steps occur in response to receiving a spoken utterance. Those steps include extracting prosodic parameters from and performing speech recognition upon that received utterance. These steps are then followed by synthesizing a nominal word from the recognized word and generating a prosodic mimic word from the synthesized nominal word and the prosodic parameters

that were extracted from that received utterance. Furthermore, there is finally an executing/dialing step that performs a function relating to that received utterance.

In the described embodiment, the purpose of performing this sequence of steps is to first generate a prompt confirming that the device has correctly recognized the spoken command (or name) and then execute a function associated with that command or name. So, it is important that the prompt (i.e., the synthesized word) be derived from the spoken utterance and that the prompt be particularly intelligible to the user. The latter is accomplished by synthesizing the recognized word using the prosodic parameters that are extracted from the very same spoken utterance.

The examiner admits that Marasek does not teach the generation of a nominal word or a system implemented on a handheld device. But we note that Marasek is missing another element of claim 1, as now amended. Marasek neither executes a command that is within the received spoken utterance nor dials a number corresponding to a name recognized within the spoken utterance. Indeed, Marasek does not even disclose executing commands in response to receiving a spoken utterance, let alone a spoken utterance that includes the command.

None of the other cited references supply this additional missing element. Lumelsky does not perform the execution/dialing step. Indeed, nowhere does Lumelsky even hint that any of the words spoken by the narrator are commands to be executed or names for which a corresponding number is to be dialed. Instead, Lumelsky's system stores the received information. Lumelsky's system is designed for a very different purpose from that to which the claimed invention is applied. Lumelsky's system generates a representation of a block of text that is more compact than the actual spoken text so that when sent it uses less bandwidth than if the spoken text was sent. So, instead of sending text from which speech will be synthesized, Lumelsky sends a phonetic representation of that text that was generated from the incoming text using a text-to-speech system, and along with that phonetic representation he also sends prosodic

parameters that were extracted by comparing speech generated from a speech synthesizer associated with the text-to-speech system with speech from a narrator reading that text. The phonetic representation when combined with the prosodic parameters will enable the receiver to synthesize spoken text that sounds like it was spoken by the narrator even though it was actually synthesized by a device. The function of Lumelsky's various storage and synthesis steps is to enable a human operator to adjust the prosodic parameters associated with the stored phonetic representation speech so that the text sounds as much like the narrator as possible when it is resynthesized by the users to whom it is sent. Thus Lumelsky's process is just concerned with refining the compact speech that is stored for later synthesis. Nowhere does he mention executing a command that is within the speech received from the narrator, or dialing a number corresponding to a name recognized within the narrator's speech, as required by the claim.

Cameron also does not disclose performing the execution/dialing step as part of the claimed sequence of actions performed in response to receiving a single spoken utterance. That sequence of actions includes: extracting prosodic parameters, performing speech recognition to generate a recognized word, synthesizing a nominal word, generating a prosodic mimic word, and "if the recognized word includes a command, executing the command on the handheld device, and if the recognized word includes a name, dialing a number corresponding to the name." Instead, Cameron uses a two stage process in which the stages take place in response to different utterances. In the first stage, which corresponds to a training/setup mode, he receives a user's speech and stores it as compressed speech data. In the second stage, he receives speech that includes a command, identifies the command, and retrieves and outputs the stored speech data corresponding to the command. Cameron summarizes his method as follows:

> 1) generating and storing compressed speech data corresponding to a user's speech received through the input transducer; 2) comparing the stored speech data; 3) resynthesizing the stored speech data for output as speech through the output transducer; 4) providing an audible user interface including a speech assistant for providing instructions in the user's language; 5) storing user-specific compressed speech data, including commands, received in response to

> prompts from the speech assistant for purposes of adapting the system to the user's speech; 6) identifying memo management commands spoken by the user, and storing and organizing compressed speech data as a function of the identified commands; and 7) identifying memo retrieval commands spoken by the user, and retrieving and outputting the stored speech data as a function of the commands. (p.4 ll. 6-18, emphasis added)

Thus, steps 1 and 5 correspond to the first stage, i.e., receiving speech from a user and storing it as compressed speech data, and steps 3 and 7 correspond to the second stage, i.e., identifying a command and outputting the speech data previously stored in the first stage. In other words, Cameron's system is a "voice assistant [that] operates as a user interface and is a collection of pre-recorded prompts and instructions that the PDA program plays according to user input." (p. 8, ll. 16-18, emphasis added) Thus nowhere does Cameron even hint that in response to receiving a spoken utterance, his system generates a prosodic mimic word derived from that utterance and executes a command or dials a number corresponding a recognized word from that same utterance. Instead, Cameron's two stage system receives and stores speech from one or more prior user utterances, and then, in response to a subsequent utterance, identifies a command in the subsequent utterance, outputs the stored speech, and executes the command.

In view of the above, Applicants believe that claim 1 is patentable over the cited references. Independent claim 9 contains limitations that are analogous to those of claim 1. Therefore, for the reasons discussed above, Applicants believe that claim 9, and dependent claims 2-4, 6-8, 11, 12, and 15 are also patentable over the cited references.

For the reasons stated above, we believe that the claims are in condition for allowance and therefore ask the Examiner to allow them to issue.

Please apply any charges not covered, or any credits, to Deposit Account No. 08-0219, under Order No. 112855.122US2 from which the undersigned is authorized to draw.

Respectfully submitted,

Dated: July 11, 2008

/Oliver Strimpel/
Oliver Strimpel
Registration No.: 56,451
Attorney for Applicant(s)

Wilmer Cutler Pickering Hale and Dorr LLP
60 State Street
Boston, Massachusetts 02109
(617) 526-6000 (telephone)
(617) 526-5000 (facsimile)